Amendments to the Specification:

Please replace the "BACKGROUND" section in specification as permitted under 37 C.F.R. §1.121(b) with the following amended "BACKGROUND" section:

[0072] BACKGROUND

[0073] Dynamic Frame Size Adjustment & Selective Reject

[0074] The widespread availability of personal computers at low cost has lead to a situation where the general public increasingly demands access to the Internet and other computer networks. A similar demand exists for wireless communications in that the public increasingly demands that cellular telephones be available at low cost with widespread coverage.

[0075] As a result of their familiarity with these two technologies, the general population now increasingly wishes to not only have access to computer networks, but also wishes to access such networks in wireless fashion as well. This is of particular concern for the users of portable computers, laptop computers, hand-held personal digital assistants (PDAs), and the like, who would prefer and indeed now expect to be able to access such networks with the same convenience they have grown accustom to when using their cellular telephones.

[0076] Unfortunately, there is still no widely available satisfactory approach for providing low cost, high speed access to the Internet and other networks using the existing wireless infrastructure which has been built at some expense to support cellular telephony. Indeed, at the present time, the users of wireless modems that operate with the existing cellular telephone network often experience a difficult time when trying to, for example, use the Internet to view web pages. The same

frustration level is felt in any situation when attempting to perform other tasks that require the transfer of relatively large amounts of data between computers.

[0077] This is at least in part due to the architecture of cellular telephone networks, which were originally designed to support voice communications, as compared to the communication protocols in use for the Internet, which were originally optimized for wireline communication. In particular, the protocols used for connecting computers over wireline networks do not lend themselves well to efficient transmission over standard wireless connections.

[0078] For example, cellular networks were originally designed to deliver voice grade services, having an information bandwidth of approximately three kilohertz (kHz). While techniques exist for communicating data over such radio channels at rate of 9600 k/bits per second (kbps), such low frequency channels do not lend themselves directly to transmitting data at rates of 28.8 kbps or even the 56.6 kbps that is now commonly available using inexpensive wireline modems. These rates are presently thought to be the minimum acceptable data rates for Internet access.

[0079] This situation is true for advanced digital wireless communication protocols as well, such as Code Division Multiple Access (CDMA). Even though such systems convert input voice information to digital signals, they were also designed to provide communication channels at voice grade bandwidth. As a result, they have been designed to use communication channels that may exhibit a bit error rate (BER) of as high as approximately one in one thousand bits in multipath fading environments. While such a bit error rate is perfectly acceptable for the transmission of voice signals, it becomes cumbersome for most data transmission environments.

[0080] Such a high bit error rate is certainly unacceptable for Internet type data transmissions. For example, the Transmission Control Protocol/Internet Protocol (TCP/IP) standard in use for Internet air transmission uses a frame size of 1480 bits. Thus, if a bit error is received in every frame, such as detected by a frame check sequence, it would appear as though every single frame might have to be re-transmitted in certain applications.

[0081] Dynamic Bandwidth Allocation

[0082] The increasing use of wireless telephones and personal computers by the general population has led to a corresponding demand for advanced telecommunication services that were once thought to only be meant for use in specialized applications. In the late 1980's, wireless voice communication such as available with cellular telephony had been the exclusive province of the businessman because of expected high subscriber costs. The same was also true for access to remotely distributed computer networks, whereby until very recently, only business people and large institutions could afford the necessary computers and wireline access equipment. As a result of the widespread availability of both technologies, the general population now increasingly wishes to not only have access to networks such as the Internet and private intranets, but also to access such networks in a wireless fashion as well. This is particularly of concern for the users of portable computers, laptop computers, hand-held personal digital assistants and the like who would prefer to access such networks without being tethered to a telephone line.

[0083] There still is no widely available satisfactory solution for providing low cost, high speed access to the Internet and other networks using existing wireless networks. This situation is most likely an artifact of several unfortunate circumstances. For one, the typical manner of providing high speed data service in

the business environment over the wireline network is not readily adaptable to the voice grade service available in most homes or offices. Such standard high speed data services do not lend themselves well to efficient transmission over standard cellular wireless handsets. Furthermore, the existing cellular network was originally designed only to deliver voice services. As a result, the emphasis in present day digital wireless communication schemes lies with voice, although certain schemes such as CDMA do provide some measure of asymmetrical behavior for the accommodation of data transmission. For example, the data rate on an IS-95 forward traffic channel can be adjusted in increments from 1.2 kilobits per second (kbps) up to 9.6 kbps for so-called Rate Set 1 and in increments from 1.8 kbps up to 14.4 kbps for Rate Set 2. On the reverse link traffic channel, however, the data rate is fixed at 4.8 kbps.

[0084] At present, the wireless modulation schemes in use continue their focus on delivering voice information with maximum data rates only in the range of 9.6 kbps being readily available. This is because the cellular switching network in most countries, including the United States, uses analog voice channels having a bandwidth from about 300 to 3600 Hertz. Such a low frequency channel does not lend itself directly to transmitting data at rates of 28.8 kilobits per second (kbps) or even the 56.6 kbps that is now commonly available using inexpensive wire line modems, and which rates are now thought to be the minimum acceptable data rates for Internet access.

[0085] Switching networks with higher speed building blocks are just now coming into use in the United States. Although certain wireline networks, called Integrated Services Digital Networks (ISDN), capable of higher speed data access have been known for a number of years, their costs have only been recently reduced to the point where they are attractive to the residential customer, even for wireline

service. Although such networks were known at the time that cellular systems were originally deployed, for the most part, there is no provision for providing ISDN-grade data services over cellular network topologies. ISDN is an inherently circuit switched protocol, and was, therefore, designed to continuously send bits in order to maintain synchronization from end node to end node to maintain a connection. Unfortunately, in wireless environments, access to channels is expensive and there is competition for them; the nature of the medium is such that they are expected to be shared. This is dissimilar to the usual wireline ISDN environment in which channels are not intended to be shared by definition.

[0086] The design of such existing systems therefore typically provides a radio channel which can accommodate maximum data rates only in the range of 14.4 kilobits per second (kbps) at best in the forward direction. Such a low data rate channel does not lend itself directly to transmitting data at rates of 28.8 or even 56.6 kbps that are now commonly available using inexpensive wire line modems, not to mention even higher rates such as the 128 kbps which are available with Integrated Services Digital Network (ISDN) type equipment. Data rates at these levels are rapidly becoming the minimum acceptable rates for activities such as browsing web pages. Other types of data networks using higher speed building blocks such as Digital Subscriber Line (xDSL) service are just now coming into use in the United States. However, their costs have only been recently reduced to the point where they are attractive to the residential customer.

[0087] Although such networks were known at the time that cellular systems were originally deployed, for the most part, there is no provision for providing higher speed ISDN- or xDSL-grade data services over cellular network topologies. Unfortunately, in wireless environments, access to channels by multiple subscribers is expensive and there is competition for them. Whether the multiple access is

provided by the traditional Frequency Division Multiple Access (FDMA) using analog modulation on a group of radio carriers, or by newer digital modulation schemes the permit sharing of a radio carrier using Time Division Multiple Access (TDMA) or Code Division Multiple Access (CDMA), the nature of the radio spectrum is that it is a medium that is expected to be shared. This is quite dissimilar to the traditional environment for data transmission, in which the wireline medium is relatively inexpensive to obtain, and is therefore not typically intended to be shared.

[0088] Other considerations are the characteristics of the data itself For example, consider that access to web pages in general is burst-oriented, with asymmetrical data rate transmission requirements. In particular, the user of a remote client computer first specifies the address of a web page to a browser program. The browser program then sends this web page address data, which is typically 100 bytes or less in length, over the network to a server computer. The server computer then responds with the content of the requested web page, which may include anywhere from 10 kilobytes to several megabytes of text, image, audio, or even video data. The user then may spend at least several seconds or even several minutes reading the content of the page before requesting that another page be downloaded. Therefore, the required forward channel data rates, that is, from the base station to the subscriber, are typically many times greater than the required reverse channel data rates.

[0089] In an office environment, the nature of most employees' computer work habits is typically to check a few web pages and then to do something else for extended period of time, such as to access locally stored data or to even stop using the computer altogether. Therefore, even though such users may expect to remain connected to the Internet or private intranet continuously during an entire day, the

actual overall nature of the need to support a required data transfer activity to and from a particular subscriber unit is actually quite sporadic.

Furthermore, prior art wireless communication systems provide a continuous bandwidth to individual subscribers. That is, in such networks, during a communication session the bandwidth available at all times is constant and has been designed, as noted above, primarily for voice grade use.

Please replace the "SUMMARY" section in specification as permitted under 37 C.F.R. §1.121(b) with the following amended "SUMMARY" section:

[0025] SUMMARY

[0026] Dynamic Frame Size Adjustment & Selective Reject

[0027] In view of the foregoing background, an object of the present invention is to more efficiently transmit digital signals in a wireless digital communication system.

[0028] This and other objects, advantages and features in accordance with the present invention are provided by a base station providing wireless communication of digital signals, with the digital signals being communicated in frames using a radio frequency channel via Code Division Multiple Access (CDMA) modulated radio signals. The base station may include a wireless transceiver for establishing a communication session over a digital communication path, and a bandwidth management module connected to the wireless transceiver for allocating a code channel within the radio frequency channel for the digital communication path to exchange digital signals during the communication session.

[0029] The bandwidth management module may also divide a current frame of digital signals into subframes to be transmitted within the code channel. The wireless transceiver transmits the subframes over the digital communication path, and receives feedback over the digital communication path on the subframes received with errors. The bandwidth management module may adjust a size of each subframe received with errors to a more efficient subframe size to be retransmitted over the digital communication path.

[0030] The present invention is particularly advantageous in environments requiring the communication of TCP/IP protocols since the number of channels

needed to carry a single data stream at burst rates of 56.6 or 128 kbps can be quite large. For example, carrying such TCP/IP frames at these data rates may require up to and including 20 channels operating at 9.6 kbps. Because the probability of at least one relatively weak channel may be significant, by optimizing the throughput of each channel separately, the base station obtains the best overall system throughput in such environments.

[0031] The more efficient subframe sizes may be based on at least one of maximum throughput and minimum transmission time. The bandwidth management module may determines a ratio of the subframes received with errors and subframes received without errors, and uses the ratio when determining the more efficient subframe sizes. The bandwidth management module may initially determines a size of each subframe within the current frame based upon a number of subframes received with errors for a previous frame.

[0032] Each subframe may include a position identifier, a data portion, an integrity check sum and a sequence number. A subframe is considered to be received with errors over the digital communications path if the integrity check sum is not correct, the sequence number is missing, or the position identifier is missing.

[0033] The at least one code channel may comprise a plurality of code channels, and the wireless transceiver transmits the plurality of subframes over the plurality of code channels. The digital signals may comprise at least one of voice and data signals.

[0034] The wireless communication of digital signals is performed with a subscriber unit over the digital communication path. The at least one radio frequency channel may comprise first and second radio frequency channels. The first radio frequency channel establishes a forward code channel between the wireless transceiver and the subscriber unit, with the plurality of subframes being

transmitted to the subscriber unit on the forward code channel. The second radio frequency channel establishes a reverse code channel between the subscriber unit and the wireless transceiver, with the feedback on the subframes received with errors being transmitted on the reverse code channel by the subscriber unit.

[0035] Another aspect of the present invention is directed to a subscriber unit for providing wireless communication of digital signals between terminal equipment connected therewith and a digital communication path, with the digital signals being communicated in frames using at least one radio frequency channel via Code Division Multiple Access (CDMA) modulated radio signals.

[0036] The subscriber unit may comprises a wireless transceiver for establishing a communication session over the digital communication path, and a bandwidth management module connected to the wireless transceiver for receiving over the digital communication path at least one allocated code channel within the at least one radio frequency channel to exchange digital signals during the communication session.

[0037] The bandwidth management module may divide a current frame of digital signals into a plurality of subframes to be transmitted within the at least one code channel. The wireless transceiver may transmit the plurality of subframes over the digital communication path, and receives feedback over the digital communication path on the subframes received with errors. The bandwidth management module adjusts a size of each subframe received with errors to a more efficient subframe size to be retransmitted over the digital communication path.

[0038] Yet another aspect of the present invention is directed to a digital communication system comprising a subscriber unit as defined above for providing wireless communication of digital signals; and a base station as also defined above

for establishing a communication session with the subscriber unit over a digital communications path.

[0039] In an alternative embodiment, the present invention is implemented via a protocol converter disposed between a physical communication layer, such as may be associated with implementing a wireless communication protocol, and a network layer, such as may be associated with implementing a network protocol.

[0040] The protocol converter first splits messages in the form of network layer frames into multiple subframes prior to formatting them for transmission. The subframes are each assigned a position number such that they may be reassembled into the proper order to reconstruct the network layer frame at the receiver end.

[0041] The protocol preferably makes use of multiple physical layer connections such as radio links as needed to transmit the subframes at an overall desired data transmission rate. When this is the case, a link sequence identifier is added to identify the order in which the subframes are sent over a given sub-channel in a link.

[0042] On the receiver side, the subframes are then reassembled into the network layer frames using the subframe position numbers, and then passed the reassembled frame up to the network layer. Thus, the receiver side includes a protocol converter that performs the inverse function.

[0043] The protocol converters at both the sender and receiver also take steps to automatically and dynamically adjust the size of the subframes based upon an observed rejected subframe rate in order to optimize overall throughput. An average rate at which frames are rejected can be determinined by counting good subframes and bad subframes. For example, at the receive end, a subframe with a bad cyclic redundancy check code (CRC) is discarded and counted as a bad subframe. By

keeping track of the sequence numbers of the good subframe received, the receiver can determine that a particular subframe sequence number, namely the frame with the sequence number between the last good frame and the next good frame is missing. The receiver then explicitly requests retransmission of the bad frame by sequence number. This so called selective reject feature of the transmission permits both the receiver and the sender to know the number of frames received in error from the tally of selective reject orders.

[0044] From the count of the number of frames sent and the number of selective reject order received, the sender then dynamically adjusts the size of later transmitted subframes. Preferably, the subframe size is adjusted based upon a formula which depends upon the ratio of the actual data transferred to the number of bits actually used to carry the transmission, including the frame overhead and re-transmissions. For example, the number of data bytes, X, in a given subframe can be adjusted according to the formula:

$$X = -H + \sqrt{(X_{current} + H_{current}) * (H / R)}$$

where H is the new frame overhead, in bytes, including any shared frame synchronization flag (7E) between frames, $X_{current}$ and $H_{current}$ are, respectively, the immediately prior values of X and H, and R is a ratio of the observed number of frames transmitted successfully to the number of frames that are not transmitted successfully.

[0045] Particularly noisy channels may be subjected to down speed procedures or error coding techniques in order to improve the bit error rate observed in a particular channel.

[0046] In order to optimize throughput on overall basis, the subframe size calculation is preferably carried out on each channel separately. Otherwise, any good channels, that is, those channels which do not experience particularly noisy

environments, might suffer down speed procedures needed to accommodate the weakest channels.

[0047] In one specific embodiment of the invention, the physical layer radio links may be implemented as 9.6 kbps channels such as can be reliably provided using CDMA cellular protocols and subchannel coding techniques.

[0048] The invention is particularly advantageous in environments such as requiring the communication of TCP/IP protocols since the number of channels needed to carry a single data stream at burst rates of 56.6 or 128 kbps can be quite large. For example, carrying such TCP/IP frames at these data rates may require up to and including 20 channels operating at 9.6 kbps. Because the probability of at least one relatively weak channel may be significant, by optimizing the throughput of each channel separately, the invention obtains the best overall system throughput in such environments. Simulations of the implementation of the invention indicate that it may be used to provide data rates such as 128 kbps with a bit error rate of 10^{-6} or better.

[0049] Dynamic Bandwidth Allocation

[0050] Prior art methodologies for transmission of data over wireless networks suffer numerous problems. As noted above, the bandwidth available for a single subscriber unit channel is typically fixed in size. However, data communications tend to be bursty in nature, often requiring a need for large amounts of bandwidth at certain times, while requiring very little amounts, or even none, at other times. These wide swings in bandwidth requirements can be very close together in time.

[0051] For example, when browsing a web site using HyperText Transfer Protocol (HTTP), the user of a web browser typically selects pages by selecting or clicking a single link to a page causing the client computer to send a small page

request packet to the web server. The request packet in the receive link direction requires very little bandwidth. However, in response to the request, the server typically delivers one or more web pages ranging in size from 10 to 100 kilobits (kB) or more to the client in the forward link direction. To receive the pages, the bandwidth requirements are much greater than to request the pages. The optimum bandwidth needed to acceptably receive the pages is rarely realized due to the inefficiency of the present wireless protocols that only offer maximum data rates of about 9600 bps under optimal conditions. This results in the server having to hold back some of the requested data until the network can "catch up" with the data delivery and also results in frustrated users having slow response and page loading times. In essence, the bandwidth to send a request is more than is needed, and the bandwidth to receive the pages is not enough to deliver the data at acceptable rates.

[0052] Another problem with prior art systems is that the time frame between when the small page request message leaves the wireless network and becomes wirebound, and when the pages of requested data enter the wireless portion of the data communications session on the return link is often quite long. This time-from-request to time-of-receipt delay is a function of how congested the network and server are during that time.

[0053] The present invention is based in part on the observation that bandwidth is being wasted during periods of time when waiting for data from the wireline network. Prior art wireless communications systems maintain the constant availability of the full bandwidth of the 9600 bps wireless connection for that entire data communication session, even though the wireless client may be waiting for return pages. This bandwidth which is effectively unused is therefore wasted because there is no way to allocate the channel resources in use for this data communication session to another session needing more bandwidth. That is, if other

concurrent wireless data communications sessions are taking place for other subscriber units, these concurrent sessions have no way in the prior art systems to take advantage of any unused bandwidth allocated to the client merely waiting for return pages, as in this example.

[0054] The present invention provides high speed data and voice service over standard wireless connections via an unique integration of ISDN protocols and existing cellular signaling such as is available with Code Division Multiple Access (CDMA) type modulated systems. The present invention achieves high data rates through more efficient allocation of access to the CDMA wireless channels. In particular, a number of subchannels are defined within a standard CDMA channel bandwidth, which is normally necessary to support the ISDN protocol, such as by assigning different codes to each subchannel. The instantaneous bandwidth needs of each on-line subscriber unit are met by dynamically allocating multiple subchannels of the RF carrier on an as needed basis for each session. For example, multiple subchannels are granted during times when the subscriber bandwidth requirements are relatively high, such as when downloading Web pages and released during times when the line content is relatively light, such as when the subscriber is reading a Web page which has been previously downloaded or is performing other tasks.

[0055] Specifically, the invention provides a scheme for determining an efficient allocation of N fixed rate data channels amongst M users. The invention addresses the problem of how to allocate these channels in the most effective manner between users competing for channel use. For example, when more users exist than channels, the invention determines a set of probabilities for which users will require channel access at which times, and assigns channel resources accordingly. The invention can also dynamically take away or deallocate channels

(i.e., bandwidth) from idle subscribers and provide or allocate these freed-up channels to subscribers requiring this bandwidth.

[0056] Channel resources are allocated according to a buffer monitoring scheme provided on forward and reverse links between a base station and multiple subscriber units. Data buffers are maintained for each connection between a base station and a subscriber unit. Each buffer is monitored over time for threshold levels of data to be transmitted in that buffer. In essence, the thresholds measure the "fullness" of buffers over time for each respective subscriber unit monitored. For each buffer, a probability is calculated that indicates how often a specific buffer for a specific subscriber will need to transmit data and how much data will be transmitted. This probability takes into account the arrival rates of data into the buffer, as well as which thresholds within the buffer are exceeded, as well as which resources in the form of channels are already allocated to the subscriber unit. Based on this probability, channel resources for data transmission can be either allocated or deallocated to subscriber units depending upon a forecasted need.

[0057] Further, subchannel assignment algorithms may be implemented to offer various levels of priority service to particular subscribers. These may be assigned based upon available ports per subscriber, expected user bandwidth, service premium payments, and so on.

[0058] In accordance with another aspect of the invention, some portion of the available bandwidth is initially allocated to establish a communication session. Once the session has been established, if a subscriber unit has no data to present for transmission, namely, if the data path remains quiescent for some period of time, the previously assigned bandwidth is deallocated. In addition, it is preferable that not all of the previously assigned bandwidth be deallocated, but rather at least some portion be kept available for use by an in-session subscriber. If the inactivity

continues for a further period of time, then even the remaining portion of the bandwidth can be deallocated from the session. A logical session connection at a network layer protocol is still maintained even if no subchannels are assigned.

[0059] In a preferred arrangement, a single subchannel is maintained for a predetermined minimum idle time for each network layer connection. This assists with more efficient management of channel setup and tear down.

Please replace the "BRIEF DESCRIPTION OF THE DRAWINGS" section in specification as permitted under 37 C.F.R. §1.121(b) with the following amended "BRIEF DESCRIPTION OF THE DRAWINGS" section:

[0060] BRIEF DESCRIPTION OF THE DRAWINGS

[0061] The foregoing and other objects, features and advantages of the invention will be apparent from the following more particular description of preferred embodiments of the invention, as illustrated in the accompanying drawings in which like reference characters refer to the same parts throughout the different views. The drawings are not necessarily to scale, emphasis instead being placed upon illustrating the principles of the invention.

[0062] Fig. 1 is a block diagram of a system in which a portable device such a laptop computer is making use of a protocol converter according to the invention to connect to a computer network over a wireless link.

[0063] Fig. 2 is a diagram depicting how network layer data frames are divided among multiple physical links or channels.

[0064] Fig. 3 is a more detailed diagram showing how network layer frames are divided into subframes by a protocol converter located at a sender.

[0065] Fig. 4 is a continuation of the diagram of Fig. 3.

[0066] Fig. 5 is a series of steps performed by a protocol converter at the sender to implement the invention.

[0067] Fig. 6 is a continuation of the diagram of Fig. 5.

[0068] Fig. 7 is a diagram of the steps performed by a protocol converter located at a receiver to implement the invention.

[0069] Fig. 8 is a diagram of one particular embodiment of a subframe according to the invention.

[0070] Fig. 9 is a chart illustrating a particular example of how 20 twenty 9.6 kbps sub-channels with various bit error rates can be used to provide a 138 kbps overall effective transfer rate.

[0071] Fig. 10 is a plot of how the effective bit error rate changes as the number of data bytes in a subframe changes.

[0072] Fig. 11 is a block diagram of a wireless communication system making use of a bandwidth management scheme according to the invention.

[0073] Fig. 12 is an Open System Interconnect (OSI) type layered protocol diagram showing where the bandwidth management scheme is implemented in terms of communication protocols.

[0074] Fig. 13 is a diagram showing how subchannels are assigned within a given radio frequency (RF) channel.

[0075] Fig. 14 is a more detailed block diagram of the elements of a subscriber unit.

[0076] Fig. 15 is a state diagram of the operations performed by a subscriber unit to request and release subchannels dynamically.

[0077] Fig. 16 is a block diagram of a portion of a base station unit necessary to service each subscriber unit.

[0078] Fig. 17 is a high level structured English description of a process performed by the base station to manage bandwidth dynamically according to the invention.

[0079] Fig. 18 is a block diagram of an example wireless communication system making use of a bandwidth management scheme according to the invention.

[0080] Fig. 19 is a diagram showing how channels are assigned within a given radio frequency (RF) channel.

[0081] Fig. 20 is a block diagram illustrating the internal components of a base station and subscriber units that provide the dynamic bandwidth allocation mechanism.

[0082] Fig. 21 illustrates the structure of the buffers used in either the base station or subscriber units.

[0083] Fig. 22 is a block diagram of an example wireless communication system making use of a bandwidth management scheme according to the invention.

[0084] Fig. 23 is a diagram showing how channels are assigned within a given radio frequency (RF) channel.

[0085] Fig. 24 is a diagram illustrating the protocol layers of a wireless communication system.

[0086] Fig. 25 illustrates the structure of session queues and data buffers used in the base station.

[0087] Fig. 26 is a buffer level diagram.

[0088] Fig. 27 is a buffer level diagram when resources are being added.

[0089] Fig. 28 is a buffer level diagram when resources are being taken away.

Please replace the "DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS" section in specification as permitted under 37 C.F.R. §1.121(b) with the following amended "DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS" section:

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

[0090] Turning attention now to the drawings more particularly, Fig. 1 is a block diagram of a system 10 for implementing high speed data communication according to the invention. The system 10 comprises a remote or subscriber unit 20, multiple bi-directional communication links 30, and a local or service provider unit 40.

[0091] The subscriber unit 20 connects to terminal equipment 22 such as a portable or laptop computer, hand held Personal Digital Assistant (PDA) or the like, via a modem 24. The modem 24 in turn provides data to a protocol converter 25, which in turn provides data to a multichannel digital transceiver 26 and antenna 27.

[0092] The modem 24 receives data from the terminal equipment 22, and together with appropriate hardware and/or software, converts it to a format suitable for transmission such as in accordance with known communication standards. For example, the modem 24 may convert data signals from the terminal equipment 22 to a wireline physical layer protocol format such as specified by the Integrated Services Digital Network (ISDN) standard at rates of 128 kbps, or the Kflex standard at rates of 56.6 kbps. At a network layer, the data provided by the modem is preferably formatted in a manner consistent with suitable network communication protocols such as TCP/IP to permit the terminal equipment 22 to connect to other computers over networks such as the Internet. This description of

the modem 24 and protocols is exemplary only and it should be understood that other protocols can be used.

[0093] The protocol converter 25 implements an intermediate protocol layer for converting the data provided by the modem 24 to a format appropriate for the multichannel transceiver 26 according to the invention, and as will be described in much grater detail below.

[0094] The multichannel digital transceiver 26 provides access to one or more physical communication links such as the illustrated radio channels 30. The physical links are preferably known wireless communication air interfaces using digital modulation techniques such as Code Division Multiple Access (CDMA) standard specified by IS-95. It should be understood that other wireless communication protocols and other types of links 30 may also be used to advantage with the invention.

[0095] The channels 30 represent one or more relatively slower communication channels, such as operating at a 9.6 kbps rate typical of voice grade communication. These communications channels may be provided by a single wide bandwidth CDMA carrier such as having a 1.25 MegaHertz bandwidth, and then providing the individual channels with unique orthogonal CDMA codes. Alternatively, the multiple channels 30 may be provided by single channel communication media such as provided by other wireless communication protocols. However, what is important is that the net effect is that the channels 30 represent multiple communication channels that may be adversely effected by significant bit error rates that are unique to each link 30.

[0096] An "error" as described herein is a bit error perceived at the higher layer such as the network layer. The invention primarily strives to improve the system level bit error rate instead of providing absolute data integrity.

[0097] On the local level, the service provider equipment 40 may, for example, be implemented at a wireless Internet Service Provider (ISP) 40-1. In this case, the equipment includes an antenna 42-1, a multichannel transceiver 44-1, a protocol converter 46-1, and other equipment 48-1 such as modems, interfaces, routers, and the like which are needed for the ISP to provide connections to the Internet 49-1.

[0098] At the ISP 40-1, the multichannel transceiver 44-1 provides functions analagous to the multichannel transceiver 26 of the subscriber unit, but in an inverse fashion. The same is true of the protocol converter 46-1, that is, it provides inverse functionality to the protocol converter 25 in the subscriber unit 20. The ISP 40-1 accepts data from the protocol converter 46-1 in the TCP/IP frame format and then communicates such data to the Internet 49-1. It should be understood that the configuration of the remaining ISP equipment 48-1 may take any number of forms such as a local area networks, multiple dial up connections, T1 carrier connection equipment, or other high speed communication links to the Internet 49-1.

[0099] Alternatively, the provider 40 may function as a radio base station in a cellular telephone system to permit a dial-up connection between the terminal equipment 22 and a server 49-2. In this instance, the base station 40-2 includes an antenna 42-2, multichannel transceiver 44-2, and protocol converter 46-2 providing one or more connections to a public switched telephone network (PSTN) 48-2, and ultimately to the server 49-2.

[0100] In addition to the illustrated implementations 40-1, 40-2, there may be various other ways of implementing the provider 40 in order to provide a connection to data processing equipment from the terminal equipment 22.

[0101] Turning attention now to the functions of the protocol converters 25 and 46, they can be thought of as intermediate layers within the context of the Open System Interconnect (OSI) model for communication. In particular, the protocol

converter provides a bandwidth management functionality 29 implemented between a physical layer such as provided by the CDMA protocol in use with the multichannel transceivers 26 and a network layer protocol such as TCP/IP providing connections between the terminal equipment 22 and the Internet 49-1 or server 49-2.

[0102] The bandwidth management functionality 29 preferably provides a number of functions in order to keep both the physical layer and network layer connections properly maintained over multiple communication links 30. For example, certain physical layer connections may expect to receive a continuous stream of synchronous data bits regardless of whether terminal equipment at either end actually has data to transmit. Such functions may also include rate adaption, bonding of multiple channels on the links, spoofing, radio channel setup and takedown. The details for implementing a protocol converter specifically for ISDN terminal equipment 22 and Code Division Multiple Access (CDMA) modulation techniques in use by the multichannel transceiver 26 are more specifically described in U.S. Pat. Nos. 6,151,332 and 6,081,536 assigned to the current as assignee of the present application, and which are hereby incorporated by reference in their entirety.

[0103] The present invention is more particularly concerned with the technique used by the protocol converters 25 and 46 for adjusting the frame size of individual channels used over each of the multiple links 30 in order to improve the effective throughput rate between a sender and a receiver in a bit error rate prone environment. It should be understood in the following discussion that the connections discussed herein are bidirectional, and that a sender may either be the subscriber unit 20 or the provider unit 40.

[0104] More specifically, the problem addressed by the present invention is shown in Fig. 2. The frame 60 as received at the receiver end should be identical to the frame 50 originating at the sender. This is despite the fact that multiple channels are used with much higher bit error rates with the received frame 60 being transmitted reliably with a bit error rate of 10.sup.-6 or better as is typically required in TCP/IP or other network layer protocols. The present invention increases the effective data throughput such that the received frames 60 are not affected by the experienced bit error rate performance of network layer connections.

[0105] It should be understood that another assumption is that the individual channels 30-1, 30-2 . . . 30-N may experience different bit error rate levels both over time and in an average sense. Although each of the channels 30 may operate quite similarly, given the statistical nature of errors, identical behavior of all of the channels 30 is not assumed. For example, a specific channel 30-3 may receive severe interference from another connection in a neighboring cell, and be capable of providing only a 10-3 whereby other channels 30 may experience very little interference.

[0106] To increase the throughput for the system 10 on a global basis, the invention also preferably optimizes the parameters of each channel 30 separately. Otherwise, a relatively good channel 30-1 might suffer down speed procedures required to accommodate a weaker channel 30-3.

[0107] It should also be understood that the number of channels 30 that may be needed to carry a single data stream such as a rate of 128 kbps at a given point in time may be relatively large. For example, up to 20 channels 30 may be assigned at a particular time in order to accommodate a desired data transfer rate. Therefore, the probability of different characteristics in any given one of the channels 30 is significantly different.

[0108] Turning attention now more particularly to Fig. 3, the operations of the protocol converter 25 or 46 at the sender will be more particularly described. As shown, the input frame 50 as received from the network layer is relatively large, such as for example 1480 bits long, in the case of a TCP/IP frame.

[0109] The input frame 50 is first divided into a set of smaller pieces 54-1, 54-2. The size of the individual pieces 54 are chosen based upon the optimum subframe size for each of the channels 30 available. For example a bandwidth management function may make only a certain number of channels 30 available at any time. A subset of the available channels 30 is selected, and then the optimum number of bits for each subframe intended to be transmitted over respective one of the channels, is then chosen. Thus, as illustrated in the figure, a given frame 54-1 may be divided into pieces associated with four channels. At a later time, there may be nine channels 30 available for a frame, with different optimum subframe sizes for the piece 54-2.

[0110] Each of the subframes 56 includes a position identifier 58a, a data portion 58b, and a trailer typically in the form of an integrity checksum such as a cyclic redundancy check (CRC) 58c. The position identifier 58a for each subframe indicates the position within the associated larger frame 50.

[0111] The subframes 56 are then further prepared for transmission on each channel 30. This may be done by adding a sequence number related to each channel at the beginning of each subframe 56. The subframe 56 is then transmitted over the associated channel 30.

[0112] Fig. 4 illustrates the operations performed at the receive side. The subframes 56 are first received on the individual channels 30. A subframe 56 is discarded as received if the CRC portion 58c is not correct.

[0113] The sequence numbers 58d of the remaining frames 56 are then

stripped off and used to determine whether any subframes 56 are missing. Missing subframes 56 can be detected by comparing the received sequence numbers 58d. If a sequence number is missing, it is assumed that the associated subframe 56 was not received properly. It should be understood that appropriate buffing of data and subframes 56 is typically required in order to properly receive the subframes 56 and determine if there are any missing sequence numbers depending upon the transmission rates, number of channels 30 and propagation delays in effect.

[0114] Upon the detection of a missing subframe 56, retransmission of the missed subframe is requested by the receiving end. At this point, the transmitting end reperforms transmission of the missing subframe.

[0115] Once all of the subframes 56 are received, the position number 58a is then used to arrange the data from the subframes 56 in the proper order to construct the output received frame 60.

[0116] At this point, also, if any piece of the large output frame 60 is still missing, such as when an end of frame command is encountered, retransmission of the corresponding subframe can also be requested at the indicated position, specifying a length for the missing piece.

[0117] Because of the use of both the position and sequence numbers, the sender and receiver know the ratio of the number of subframes received with errors to the number of frames received without errors. Also, the receiver and sender know the average subframe length for each channel. The optimum subframe size can thus be determined for each channel from these parameters as will be described more fully below.

[0118] Fig. 5 is a more detailed flow diagram of a set of operations performed by the sender in order to implement the invention. In a first state 100, the frame 50 is obtained from an upper communication layer such as the network layer. In a next

state 102, the sender computes an optimum subframe size from past observations of frame error rates on the individual channels 30, preferably calculating an optimum subframe size for all communication and channels available.

[0119] In a next state 104, the network layer frame 50 is divided into an appropriate number of subframes according to the optimum size for each associated channel available. This division is also based upon the available channel estimated throughput. A list of subframes is then created.

[0120] In a next state 106, a position identifier and a cyclic redundancy check (CRC) code are added to each subframe. The position identifier is an offset within the large frame 50 as described above, to allow correct positioning of the subframe when reconstructing the frame 50 at the receive end. In a next state 108, an appropriate channel 30 is associated with each subframe depending upon the subframe size and transmit queue depth, if multiple channels are available.

[0121] Upon receipt of a retransmission request of a subframe missing at the receiver, a state 110 is entered in which an optimum subframe size is computed from the observed frame averages for the available communications channels 30. The subframe list is then used to requeue the subframe for retransmission in state 112. Processing then continues at state 108 for retransmission of the missing subframe.

[0122] Fig. 6 shows the remainder of the steps performed at the sender. In a state 114, a channel related sequence number is added to each subframe. In a next state 116, subframe separators such as flags in the form "7E" are inserted into the subframes. In addition, any zero insertion such as setting data bits to a 1 after a sequence of five zeros is performed. Other synchronization and separation and coding techniques may require that bits be inserted into the subframes at this point. For example, if a given channel 30 may make use of convolutional coding as

specified by the IS-95 standard.

[0123] In a state 118, the subframes are sent on the available channels 30. Non-data frames such as logical start, logical end and other control frames may be inserted at this point as well. In a final state 120, the sender operates on any subframe retransmission requests or positive acknowledgments of a large frame being received correctly. Another frame transmission may be indicated, for example, at this point before completion of a frame in transit.

[0124] Fig. 7 shows a detailed sequence of steps performed at the receiver. In a first state 200, the subframes are received. Any subframe with a good CRC is passed to the next following state 202. Any other received data entity with a bad CRC is discarded.

[0125] Continuing with state 202, the receiver determines any missing sequence numbers. The receiver then requests retransmission of a subframe for the missing pieces based upon sequence number by sending back a retransmission request to the sender.

[0126] In a next state 204, from the position identifier and the known length of each original frame 50, the receiver attempts to rebuild the original frame 50. In state 206, if any pieces of the frame 50 are still missing after the retransmission requests are all processed, accommodating the fact that a retransmission request itself may be lost, the receiver requests the missing portion of the large frame 50 by position and size. In state 208, once the frame 50 is completely received, a positive acknowledgment is returned back to the sender.

[0127] Fig. 8 is a diagram illustrating the format of a typical subframe 56. The fields include a data/command field, a large frame sequence number field of two bits, a position (character) offset field of the subframe into the large frame, a

channel sequence number, the data, a CRC field, and a shared subframe interframe flag. The data/command indicator and large frame sequence number field may each be comprised, for example, of one bit. The position offset of the subframe into the associated large frame may be 11 bits long. The channel sequence number may be 3 bits long. The data field varies from 0 to 2048 bits long, the CRC field may be 12 bits, and the flag may be the standard hex value "7E" of 8 bits.

[0128] Returning to Fig. 5 briefly, as mentioned above, an optimum size is computed in state 102, given a frame error ratio, in order to optimizes the frame size. The objective is to improve the perceived bit error rate, assuming that a single bit error will destroy the integrity of a large frame maximizing the efficiency of a given channel 30. The efficiency is the ratio of the actual data bits versus all the data bits transmitted, including protocol elements such as CRC, zero insertions, frame separators and other overhead bits. Another objective is to extend the optimum efficiency in a multichannel environment where each channel may have a degree of efficiency different from the other channels.

[0129] The actual measurement of the number and position of bits in error is impractical and/or time consuming in most real systems. A single bit in error destroys frame integrity, but one or more bits in error in sequence most likely produced the same damage as a single bit. Conversely, a single bit in error in the middle of a synchronization flag destroys two frames. The number of bits in a frame, therefore, cannot be determined exactly without knowing the content of the frame, due to zero insertion.

[0130] However, one practical measurement available is R, the ratio of received good frames to received bad frames. By definition, a frame is in error because of a frame integrity destroying the event. Such an event can be a single bit in error or cluster of bits in error within a frame boundary. Regardless of how the

error occurs, the optimum sub-frame size can be determined using the following equations, given information about the frame error rate. Consider first the following definitions:

[0131] G number of good bits received, on average, before a bit is received in error H frame overhead, in bytes, including any shared frame synchronization flag (7E) between frames;

[0132] X number of data bytes in a frame;

[0133] B total number of bytes in a frame, including data plus overhead;

[0134] N number of original data frames, i.e., the number of frames generated by the sender;

[0135] F total number of frames transmitted, including bad frames and re-transmitted frames; and the frame error ratio, R, can be defined as:

$$R=F_{RB}/F_{RG} \qquad (1)$$

[0136] where F_{RB} is the number of frames observed to be received in error and F_{RG} is the observed number of frames correctly received at the receive end.

[0137] After attempting the transmission of N frames, some of them are received correctly, and some will have been received in error. Some of the latter, in turn, will be re-transmitted, and require still further re-transmission. In general,

$$F=N+N*R+(N*R)*R+(N*R*R)*R+ \qquad (2)$$

$$F=N*(1+R^2+R^3+\ldots) \qquad (3)$$

$$F=N/(1-R) \qquad (4)$$

[0138] A normalized efficiency, F_n, can be defined for N=1 as:

$$F_n=1/(1-R) \qquad (5)$$

[0139] An efficiency of transmission, K, can in turn be defined as the ratio of data bytes to the total number of data bytes required to transmit the original data, including re-transmissions and frame overhead:

$$K = X/(B * F_n) \tag{6}$$

$$= \frac{X}{(X+H)/(1-R)} \tag{7}$$

$$= \frac{X*(1-R)}{X+H} \tag{8}$$

$$= \frac{X*(1-((X+H)*8/G))}{(X+H)} \tag{9}$$

$$= \frac{X - X*8*(X+H)/G}{(X+H)} \tag{10}$$

$$= \frac{X}{(X+H)} - \frac{X*8*(Z+H)/G}{(X+H)} \tag{11}$$

$$= \frac{X}{(X+H)} - \frac{8*X}{G} \tag{12}$$

[0140] In order to optimize the efficiency of transmission, K, it is necessary to find the maximum of the above function. This can be done by setting the derivative of K to zero:

$$= \frac{dK}{dK} = \frac{1}{(X+H)} * \frac{d}{dX}(X) - \frac{X}{(X+H)^2} * \frac{d}{dX}(X+H) - \frac{8}{G} + \frac{d}{dX}(X) \tag{13}$$

or

$$0 = \frac{1}{(X+H)} - \frac{X}{(X+H)^2} - \frac{8}{G} \tag{14}$$

[0141] which, when multiplying by $(X+H)^2$ becomes:

$$X + H - X = (8/G)*(X+H)^2 \tag{15}$$

[0142] which can then be solved as

$$X + H = \sqrt{G + H/8} \tag{16}$$

or

$$(X+H)^2 = G * H / 8 \tag{17}$$

[0143] This last equation opens the possibility of implementing an algorithm that optimizes the frame size knowing the bit error rate for a specific channel. Consider that the sender knows R (by counting the number of re-transmission requests), and also knows the current X and H used to pack frames. Redefining G as the average distance, in bits, between frame integrity destroying events, G can be derived as:

$$G = (X_{current} + H_{current}) * 8 / R \tag{18}$$

[0144] By substituting this expression for G into equation (17) above of the optimization of (X+H),

$$(X+H)^2 = \frac{(X_{current} + H_{current}) * 8}{R} * \frac{H}{8} \tag{19}$$

$$X = -H + \sqrt{H_{current} + H_{current} * H / R} \tag{20}$$

[0145] This last equation is relatively straightforward to implement. The system 10 need only keep a filtered average of the number of frames transmitted successfully and the number of frames that did not go across the link. The number of data bytes in the new sub-frames are then adjusted according to the formula.

[0146] For practical purposes, there is no need for extreme accuracy in the optimization calculation since there is no guarantee that R remains constant over time. Actually, the purpose is to adapt to changes in the value of R, while still providing a frame size that optimizes the effective throughput during a period of time when only the measurement of an average for R makes sense.

[0147] Fig. 9 is a chart showing the results of modeling the system choosing

optimum frame lengths as described. The model illustrated that using a mixed set of channels 30, such as with two 9.6 kbps channels having an error every 50 bits, 5 channels having a bit error every 500 bits, and 13 channels having a bit error every 5000 bits, the system 10 can carry 138 kpbs data rate load with a perceived error rate of 10^{-6} or better.

[0148] Fig. 10 is a set ~~sets~~ of curves of the overall effective bit error rate for subchannels operating at 9.6 kbps, assuming bit error rates of 1 bit in every 100, 300, 900, 2700, 8100 and 24300 bits, respectively. Note that the peaks of the curves change depending upon the number of data bytes in a frame as well as the bit error rate.

[0149] Referring to Figures 11-17, these figures describe a dynamic bandwidth allocation process to transmit a wireless protocol across a CDMA radio link.

[0150] More specifically, Fig. 11 is a block diagram of a system 1100 for providing high speed data and voice service over a wireless connection by seamlessly integrating a digital data protocol such as, for example, Integrated Services Digital Network (ISDN) with a digitally modulated wireless service such as Code Division Multiple Access (CDMA).

[0151] The system 1100 consists of two different types of components, including subscriber units 1101, 1102 and base stations 1170. Both types of these components 1101 and 1170 cooperate to provide the functions necessary in order to achieve the desired implementation of the invention. The subscriber unit 1101 provides wireless data services to a portable computing device 1110 such as a laptop computer, portable computer, personal digital assistant (PDA) or the like. The base station 1170 cooperates with the subscriber unit 1101 to permit the transmission of data between the portable computing device 1110 and other devices such as those

connected to the Public Switched Telephone Network (PSTN) 1180.

[0152] More particularly, data and/or voice services are also provided by the subscriber unit 1101 to the portable computer 1110 as well as one or more other devices such as telephones 1112-1, 1112-2 (collectively referred to herein as telephones 1112. (The telephones 1112 themselves may in turn be connected to other modems and computers which are not shown in Fig. 11). In the usual parlance of ISDN, the portable computer 1110 and telephones 1112 are referred to as terminal equipment (TE). The subscriber unit 1101 provides the functions referred to as a network termination type 1 (NT-1). The illustrated subscriber unit 1101 is in particular meant to operate with a so-called basic rate interface (BRI) type ISDN connection that provides two bearer or "B" channels and a single data or "D" channel with the usual designation being 2B+D.

[0153] The subscriber unit 1101 itself consists of an ISDN modem 1120, a device referred to herein as the protocol converter 1130 that performs the various functions according to the invention including spoofing 1132 and bandwidth management 1134, a CDMA transceiver 1140, and subscriber unit antenna 1150. The various components of the subscriber unit 1101 may be realized in discrete devices or as an integrated unit. For example, an existing conventional ISDN modem 1120 such as is readily available from any number of manufacturers may be used together with existing CDMA transceivers 1140. In this case, the unique functions are provided entirely by the protocol converter 1130 which may be sold as a separate device. Alternatively, the ISDN modem 1120, protocol converter 1130, and CDMA transceiver 1140 may be integrated as a complete unit and sold as a single subscriber unit device 1101.

[0154] The ISDN modem 1120 converts data and voice signals between the terminal equipment 1110 and 1112 to format required by the standard ISDN "U"

interface. The U interface is a reference point in ISDN systems that designates a point of the connection between the network termination (NT) and the telephone company.

[0155] The protocol converter 1130 performs spoofing 1132 and basic bandwidth management 1134 functions, which will be described in greater detail below. In general, spoofing 1132 consists of insuring that the subscriber unit 1101 appears to the terminal equipment 1110, 1112 that is connected to the public switched telephone network 1180 on the other side of the base station 1170 at all times.

[0156] The bandwidth management function 1134 is responsible for allocating and deallocating CDMA radio channels 1160 as required. Bandwidth management also includes the dynamic management of the bandwidth allocated to a given session by dynamically assigning sub-portions of the CDMA channels 1160 in a manner which is more fully described below.

[0157] The CDMA transceiver 1140 accepts the data from the protocol converter 1130 and reformats this data in appropriate form for transmission through a subscriber unit antenna 1150 over CDMA radio link 1160-1. The CDMA transceiver 1140 may operate over only a single 1.25 MHZ radio frequency channel or, alternatively, in a preferred embodiment, may be tunable over multiple allocatable radio frequency channels.

[0158] CDMA signal transmissions are then received at the base station and processed by the base station equipment 1170. The base station equipment 1170 typically consists of multichannel antennas 1171, multiple CDMA transceivers 1172, and a bandwidth management functionality 1174. Bandwidth management controls the allocation of CDMA radio channels 1160 and subchannels. The base station 1170 then couples the demodulated radio signals to the Public Switch

Telephone Network (PSTN) 1180 in a manner which is well known in the art. For example, the base station 1170 may communicate with the PSTN 1180 over any number of different efficient communication protocols such as primary rate ISDN, or other LAPD based protocols such as IS-634 or V5.2.

[0159] It should also be understood that data signals travel bidirectionally across the CDMA radio channels 1160, i.e., data signals originate at the portable computer 1110 are coupled to the PSTN 1180, and data signals received from the PSTN 1180 are coupled to the portable computer 1110.

[0160] Other types of subscriber units such as unit 1102 may be used to provide higher speed data services. Such subscriber units 1102 typically provide a service referred to as nB+D type service that may use a so-called Primary Rate Interface (PRI) type protocol to communicate with the terminal equipment 1110, 1112. These units provide a higher speed service such as 512 kbps across the U interface. Operation of the protocol converter 1130 and CDMA transceiver 1140 are similar for the nB+D type subscriber unit 1102 as previously described for subscriber unit 1101, with the understanding that the number of radio links 1160 to support subscriber unit 1102 are greater in number or each have a greater bandwidth.

[0161] Turning attention now to Fig. 12, the invention may be described in the context of an Open Systems Interconnect multilayer protocol diagram. The three protocol stacks 1220, 1230, and 1240 are for the ISDN modem 1120, protocol converter 1130, and base station 1170, respectively.

[0162] The protocol stack 1220 used by the ISDN modem 1120 is conventional for ISDN communications and includes, on the terminal equipment side, the analog to digital conversion (and digital to analog conversion) 1221 and digital data formatting 1222 at layer one, and an applications layer 1223 at layer two. On the U

interface side, the protocol functions include Basic Rate Interface (BRI) such as according to standard 1.430 at layer one, a LAPD protocol stack at layer two, such as specified by standard Q.921, and higher level network layer protocols such as Q.931 or X.227 and high level end to end signaling 1228 required to establish network level sessions between modes.

[0163] The lower layers of the protocol stack 1220 aggregate two bearer (B) channels to achieve a single 128 kilobits per second (kbps) data rate in a manner which is well known in the art. Similar functionality can be provided in a primary rate interface, such as used by subscriber unit 1102, to aggregate multiple B channels to achieve up to 512 kbps data rate over the U interface.

[0164] The protocol stack 1230 associated with the protocol converter 1130 consists of a layer one basic rate interface 1231 and a layer two LAPD interface 1232 on the U interface side, to match the corresponding layers of the ISDN modem stack 1220.

[0165] At the next higher layer, usually referred to as the network layer, a bandwidth management functionality 1235 spans both the U interface side and the CDMA radio link side of the protocol converter stack 1230. On the CDMA radio link side 1160, the protocol depends upon the type of CDMA radio communication in use. An efficient wireless protocol referred to herein as EW[x] 1234, encapsulates the layer one 1231 and layer two 1232 ISDN protocol stacks in such a manner that the terminal equipment 1110 may be disconnected from one or more CDMA radio channels without interrupting a higher network layer session.

[0166] The base station 1170 contains the matching CDMA 1241 and EW[x] 1242 protocols as well as bandwidth management 1243. On the PSTN side, the protocols may convert back to basic rate interface 1244 and LAPD 1245 or may also include higher level network layer protocols as Q.931 or V5.2 246.

[0167] Call processing functionality 1247 allows the network layer to set up and tear down channels and provide other processing required to support end to end session connections between nodes as is known in the art.

[0168] The spoofing function 1132 performed by the EW[x] protocol 1234 includes the necessary functions to keep the U interface for the ISDN connection properly maintained, even in the absence of a CDMA radio link 1160 being available. This is necessary because ISDN, being a protocol originally developed for wire line connections, expects to send a continuous stream of synchronous data bits regardless of whether the terminal equipment at either end actually has any data to transmit. Without the spoofing function 1132, radio links 1160 of sufficient bandwidth to support at least a 192 kbps data rate would be required throughout the duration of an end to end network layer session, whether or not data is actually presented.

[0169] EW[x] 1234 therefore involves having the CDMA transceiver 1140 loop back these synchronous data bits over the ISDN communication path to spoof the terminal equipment 1110, 1112 into believing that a sufficiently wide wireless communication link 1160 is continuously available. However, only when there is actually data present from the terminal equipment to the wireless transceiver 1140 is wireless bandwidth allocated. Therefore, unlike the prior art, the network layer need not allocate the assigned wireless bandwidth for the entirety of the communications session. That is, when data is not being presented upon the terminal equipment to the network equipment, the bandwidth management function 1235 deallocates initially assigned radio channel bandwidth 1160 and makes it available for another transceiver and another subscriber unit 1101.

[0170] In order to better understand how bandwidth management 1235 and 1243 accomplish the dynamic allocation of radio bandwidth, turn attention now to

Fig. 3. This figure illustrates one possible frequency plan for the wireless links 1160 according to the invention. In particular, a typical transceiver 1170 can be tuned on command to any 1.25 MHZ channel within a much larger bandwidth, such as up to 30 MHZ. In the case of location in an existing cellular radio frequency bands, these bandwidths are typically made available in the range of from 800 to 900 MHZ. For personal communication systems (PCS) type wireless systems, the bandwidth is typically allocated in the range from about 1.8 to 2.0 GigaHertz (GHz). In addition, there are typically two matching bands active simultaneously, separated by a guard band, such as 80 MHZ; the two matching bands form forward and reverse full duplex link.

[0171] Each of the CDMA transceivers, such as transceiver 1140 in the subscriber unit 1101 and transceivers 1172 in the base station 1170, are capable of being tuned at any given point in time to a given 1.25 MHZ radio frequency channel. It is generally understood that such 1.25 MHZ radio frequency carrier provides, at best, a total equivalent of about 500 to 600 kbps maximum data rate transmission within acceptable bit error rate limitations.

[0172] In the prior art, it was thus generally understood that in order to support an ISDN type like connection which may contain information at a rate of 128 kbps that, at best, only about (500 kbps/128 kbps) or only 3 ISDN subscriber units could be supported at best.

[0173] In contrast to this, the present invention subdivides the available approximately 500 to 600 kbps bandwidth into a relatively large number of subchannels. In the illustrated example, the bandwidth is divided into 64 subchannels, each providing an 8 kbps data rate. A given subchannel is physically implemented by encoding a transmission with one of a number of different assignable pseudorandom codes. For example, the 64 subchannels may be defined

within a single CDMA RF carrier by using a different orthogonal Walsh codes for each defined subchannel.

[0174] The basic idea behind the invention is to allocate the subchannels only as needed. For example, multiple subchannels are granted during times when a particular ISDN subscriber unit 1101 is requesting that large amounts of data be transferred. These subchannels are released during times when the subscriber unit 1101 is relatively lightly loaded.

[0175] Before discussing how the subchannels are preferably allocated and deallocated, it will help to understand a typical subscriber unit 1101 in greater detail. Turning attention now to FIG. 14, it can be seen that an exemplary protocol converter 1130 consists of a microcontroller 1410, reverse link processing 1420, and forward link processing 1430. The reverse link processing 1420 further includes ISDN reverse spoofer 1422, voice data detector 1423, voice decoder 1424, data handler 1426, and channel multiplexer 1428. The forward link processing 1430 contains analogous functions operating in the reverse direction, including a channel multiplexer 1438, voice data detector 1433, voice decoder 1434, data handler 1436, and ISDN forward spoofer 1432.

[0176] In operation, the reverse link 1420 first accepts channel data from the ISDN modem 1120 over the U interface and forwards it to the ISDN reverse spoofer 1432. Any repeating, redundant "echo" bits are removed from data received and, once extracted, sent to the forward spoofer 1432. The remaining layer three and higher level bits are thus information that needs to be send over a wireless link.

[0177] This extracted data is sent to the voice decoder 1424 or data handler 1426, depending upon the type of data being processed.

[0178] Any D channel data from the ISDN modem 1120 is sent directly to voice data detection 1423 for insertion on the D channel inputs to the channel

multiplexer 1428. The voice data detection circuit 1423 determines the content of the D channels by analyzing commands received on the D channel.

[0179] D channel commands may also be interpreted to control a class of wireless services provided. For example, the controller 1410 may store a customer parameter table that contains information about the customers desired class of service which may include parameters such as maximum data rate and the like. Appropriate commands are thus sent to the channel multiplexer 1428 to request one or more required subchannels over the radio links 1160 for communication. Then, depending upon whether the information is voice or data, either the voice decoder 1424 or data handler 1426 begins feeding data inputs to the channel multiplexer 1428.

[0180] The channel multiplexer 1428 may make further use of control signals provided by the voice data detection circuits 1423, depending upon whether the information is voice or data.

[0181] In addition, the CPU controller 1410, operating in connection with the channel multiplexer 1428, assists in providing the necessary implementation of the EW[x] protocol 1234 between the subscriber unit 1101 and the base station 1170. For example, subchannel requests, channel setup, and channel tear down commands are sent via commands placed on the wireless control channel 1440. These commands are intercepted by the equivalent functionality in the base station 1170 to cause the proper allocation of subchannels to particular network layer sessions.

[0182] The data handler 1426 provides an estimate of the data rate required to the CPU controller 1410 so that appropriate commands can be sent over the control channel 1440 to allocate an appropriate number of subchannels. The data handler 1426 may also perform packet assembly and buffering of the layer three

data into the appropriate format for transmission.

[0183] The forward link 1430 operates in analogous fashion. In particular, signals are first received from the channels 1160 by the channel multiplexer 1438. In response to receiving information on the control channels 1440, control information is routed to the voice data detection circuit 1433. Upon a determination that the received information contains data, the received bits are routed to the data handler 1436. Alternatively, the information is voice information, and routed to the voice decoder 1434.

[0184] Voice and data information are then sent to the ISDN forward spoofer 1432 for construction into proper ISDN protocol format. This assembly of information is coordinated with the receipt of echo bits from the ISDN reverse spoofer 1422 to maintain the proper expected synchronization on the U interface with the ISDN modem 1120.

[0185] It can now be seen how a network layer communication session may be maintained even though wireless bandwidth initially allocated for transmission is reassigned to other uses when there is no information to transmit. In particular, the reverse 1422 and forward 1432 spoofers cooperate to loop back non-information bearing signals, such as flag patterns, sync bits, and other necessary information, so as to spoof the data terminal equipment connected to the ISDN modem 1120 into continuing to operate as though the allocated wireless path over the CDMA transceiver 1150 is continuously available.

[0186] Therefore, unless there is an actual need to transmit information from the terminal equipment being presented to the channel multiplexers 1428, or actual information being received from the channel multiplexers 1438, the invention may deallocate initially assigned subchannel, thus making them available for another subscriber unit 1101 of the wireless system 1100.

[0187] The CPU controller 1410 may also perform additional functions to implement the EW[x] protocol 1234, including error correction, packet buffering, and bit error rate measurement.

[0188] The functions necessary to implement bandwidth management 1235 in the subscriber unit 1101 are carried out in connection with the EW[x] protocol typically by the CPU controller 1410 operating in cooperation with the channel multiplexers 1428, 1438, and data handlers 1420, 1436. In general, bandwidth assignments are made for each network layer session based upon measured short term data rate needs. One or more subchannels are then assigned based upon these measurements and other parameters such as amount of data in queue or priority of service as assigned by the service provider. In addition, when a given session is idle, a connection is preferably still maintained end to end, although with a minimum number of, such as a single subchannel being assigned. For example, this single subchannel may eventually be dropped after a predetermined minimum idle time is observed.

[0189] Fig. 15 is a detailed view of the process by which a subscriber unit 1101 may request subchannel allocations from the base station 1170 according to the invention. In a first state 1502, the process is in an idle state. At some point, data becomes ready to transmit and state 1504 is entered, where the fact that data is ready to be transmitted may be detected by an input data buffer in the data handler 1426 indicated that there is data ready.

[0190] In state 1504, a request is made, such as via a control channel 1440 for the allocation of a subchannel to subscriber unit 1101. If a subchannel is not immediately available, a pacing state 1506 may be entered in which the subscriber unit simply waits and queues its request for a subchannel to be assigned.

[0191] Eventually, a subchannel is granted by the base station and the

process continues to state 1508. In this state, data transfer may then begin using the single assigned subchannel. The process will continue in this state as long as the single subchannel is sufficient for maintaining the required data transfer and/or is being utilized. However, if the input buffer should become empty, such as notified by the data handler 1426, then the process will proceed to a state 1510. In this state 1510, the subchannel will remain assigned in the event that data traffic again resumes. In this case, such as when the input buffer begins to once again become full and data is again ready to transmit, then the process returns to state 1508. However, from state 1510 should a low traffic timer expire, then the process proceeds to state 1512 in which the single subchannel is released. The process then returns to the idle state 1502. In state 1512, if a queue request is pending from states 1506 or 1516, the subchannel is used to satisfy such request instead of releasing it.

[0192] Returning to state 1508, if instead the contents of the input buffer are beginning to fill at a rate which exceeds a predetermined threshold indicating that the single subchannel is insufficient to maintain the necessary data flow, then a state 1514 is entered in which more subchannels are requested. A subchannel request message is again sent over the control channel 1440 or through a subchannel already allocated. If additional subchannels are not immediately available, then a pacing state 1516 may be entered and the request may be retried by returning to state 1514 and 1516 as required. Eventually, an additional subchannel will be granted and processing can return to state 1508.

[0193] With the additional subchannels being now available, the processing continues to state 1518 where data transfer may be made on a multiple N of the subchannels. This may be done at the same time through a channel bonding functional or other mechanism for allocating the incoming data among the N

subchannels. As the input buffer contents reduced below an empty threshold, then a waiting state 1520 may be entered.

[0194] If, however, a buffer filling rate is exceeded, then state 1514 may be entered in which more subchannels are again requested.

[0195] In state 1520, if a high traffic timer has expired, then one or more of the additional subchannels are released in state 1522 and the process returns to state 1508.

[0196] Fig. 16 is a block diagram of the components of the base station equipment 1170 of the system 1100. These components perform analogous functions to those as already described in detail in Fig. 14 for the subscriber unit 1101. It should be understood that a forward link 1620 and reverse link 1630 are required for each subscriber unit 1101 or 1102 needing to be supported by the base station 1170.

[0197] The base station forward link 1620 functions analogously to the reverse link 1420 in the subscriber unit 1100, including a subchannel inverse multiplexer 1622, voice data detection 1623, voice decoder 1624, data handler 1626, and ISDN spoofer 1622, with the understanding that the data is traveling in the opposite direction in the base station 1170. Similarly, the base station reverse link 1630 includes components analogous to those in the subscriber forward link 1430, including an ISDN spoofer 1632, voice data detection 1633, voice decoder 1634, data handler 1636, and subchannel multiplexer 1638. The base station 1170 also requires a CPU controller 1610.

[0198] One difference between the operation of the base station 1170 and the subscriber unit 1101 is in the implementation of the bandwidth management functionality 1243. This may be implemented in the CPU controller 1610 or in another process in the base station 1170.

[0199] A high level description of a software process performed by dynamic channel allocation portion 1650 of the bandwidth management 1243 is contained in Fig. 17. This process includes a main program 1710, which is continuously executed, and includes processing port requests, processing bandwidth release, and processing bandwidth requests, and then locating and tearing down unused subchannels.

[0200] The processing of port requests is more particularly detailed in a code module 1720. These include upon receiving a port request, and reserving a subchannel for the new connection, preferably chosen from the least utilized section of the radio frequency bandwidth. Once the reservation is made, an RF channel frequency and code assignment are returned to the subscriber unit 1101 and a table of subchannel allocations is updated. Otherwise, if subchannels are not available, then the port request is added to a queue of port requests. An expected waiting time may be estimated upon the number of pending port requests and priorities, and an appropriate wait message can be returned to the requesting subscriber unit 1101.

[0201] In a bandwidth release module 1730, the channel bonding function executing in the multiplexer 1622 in the forward link is notified of the need to release a subchannel. The frequency and code are then returned to an available pool of subchannels and a radio record is updated.

[0202] The following bandwidth request module 1740 may include electing the request having the highest priority with lowest bandwidth utilization. Next, a list of available subchannels is analyzed for determining the greatest available number. Finally, subchannels are assigned based upon need, priority, and availability. A channel bandwidth bonding function is notified within the subchannel multiplexer 1622 and the radio record which maintains which subchannels are assigned to which connections is updated.

[0203] In the bandwidth on demand algorithm, probability theory may

typically be employed to manage the number of connections or available ports, and the spectrum needed to maintain expected throughput size and frequency of subchannel assignments. There may also be provisions for priority service based upon subscribers who have paid a premium for their service.

[0204] It should be understood, for example, that in the case of a supporting 128 kbps ISDN subscriber unit 1101 that even more than 16x8 kbps subchannels may be allocated at a given time. In particular, one may allow a larger number, such as 20 subchannels, to be allocated to compensate for delay and reaction in assigning subchannels. This also permits dealing with bursts of data in a more efficient fashion such as typically experienced during the downloading of Web pages.

[0205] In addition, voice traffic may be prioritized as against data traffic. For example, if a voice call is detected, at least one subchannel may be active at all times and allocated exclusively to the voice transfer. In that way, voice calls blocking probability will be minimized.

[0206] While this invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention as defined by the appended claims. For example, instead of ISDN, other wireline digital protocols may be encapsulated by the EW[x] protocol, such as xDSL, Ethernet, and X.25, and therefore may advantageously use the dynamic wireless subchannel assignment scheme described herein. Those skilled in the art will recognize or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described specifically herein. Such equivalents are intended to be encompassed in the scope of the claims.

[0207] Referring to Figures 18-21, these figures describe a dynamic bandwidth allocation process for multiple access communications using a buffer urgency factor.

[0208] More specifically, Fig. 18 is a block diagram of a system 100 for providing high speed data service over a wireless connection by seamlessly integrating a digital data protocol such as, for example, Integrated Services Digital Network (ISDN) with a digitally modulated wireless service such as Code Division Multiple Access (CDMA).

[0209] The system 1800 consists of two different types of components, including subscriber units 1801-1, 1801-2, and 1801-3 (collectively subscribers 1801) as well as one or more base stations 1804 to provide the functions necessary in order to achieve the desired implementation of the invention. The subscriber units 1801 provide wireless data and/or voice services and can connect devices such as, for example, laptop computers, portable computers, personal digital assistants (PDAs) or the like through base station 1804 to a network 1805 which can be a Public Switched Telephone Network (PSTN), a packet switched computer network, or other data network such as the Internet or a private intranet. The base station 1804 may communicate with the network 1805 over any number of different efficient communication protocols such as primary rate ISDN, or other LAPD based protocols such as IS-634 or V5.2, or even TCP/IP if network 1805 is an Ethernet network such as the Internet. The subscriber units 1801 may be mobile in nature and may travel from one location to another while communicating with base station 1804.

[0210] Fig. 18 illustrates one base station 1804 and three mobile subscriber units 1801 by way of example only and for ease of description of the invention. The invention is applicable to systems in which there are typically many more subscriber units communicating with one or more base stations.

[0211] It is also to be understood by those skilled in the art that Fig. 18 may be a standard cellular type communication system such as a CDMA, TDMA, GSM or other system in which the radio channels are assigned to carry between the base stations 1804 and subscriber units 1801. This invention, however, applies more particularly to non-voice transmissions, and preferably to digital data transmissions of varying bandwidths. Thus, in a preferred embodiment, Fig. 18 is a CDMA-like system, using code division multiplexing principles for the air interface. However, it is also to be understood that the invention is not limited to using standardized CDMA protocols such as IS-95, or the newer emerging CDMA protocol referred to as IS-95B. The invention is also applicable to other multiple access techniques.

[0212] In order to provide data and voice communications between the subscriber units 1801 and base station 1804, wireless transmission of data over a limited number of radio channel resources is provided via forward communication channels 1810-a through 1810-c, and reverse communication channels 1811-a through 1811-c. The invention provides dynamic bandwidth management of these limited channel resources on an as needed basis for each subscriber unit 1801. It should also be understood that data signals travel bidirectionally across the CDMA radio channels 1810 and 1811, i.e., data signals originating at the subscriber units 1801 are coupled to the network 1805, and data signals received from the network 1805 are coupled to the subscriber units 1801.

[0213] Fig. 19 provides an example of how dynamic allocation of radio bandwidth may take place in an example system 1800. First a typical transceiver within a subscriber unit 1801 or the base station 1804 can be tuned on command to any 1.25 MegaHertz (MHZ) channel within a much larger bandwidth, such as up to 30 MHZ in the case of the radio spectrum allocated to cellular telephony; this bandwidth is typically made available in the range of from 800 to 900 MHZ in the

United States. For PCS type wireless systems, a 5 or 10 MHZ bandwidth is typically allocated in the range from about 1.8 to 2.0 GigaHertz (GHz). In addition, there are typically two matching bands active simultaneously, separated by a guard band, such as 80 MHZ; the two matching bands form a forward and reverse full duplex link between the base station 1804 and the subscriber units 1801.

[0214] For example, within the subscriber unit 1801 and the base station 1804, transmission processors (i.e., transceivers) are capable of being tuned at any given point in time to a given 1.25 MHZ radio frequency channel. It is generally understood that such 1.25 MHZ radio frequency carrier provides, at best, a total equivalent of about a 500 to 600 kbps maximum data rate transmission speed within acceptable bit error rate limitations.

[0215] In the prior art, it was thus generally understood that in order to support an ISDN type like connection which may contain information at a rate of 128 kbps that, at best, only about (500 kbps/128 kbps) or only three (3) ISDN subscriber units could be supported at best.

[0216] In contrast to this, the present invention subdivides the available approximately 500 to 600 kbps data rate among a relatively large number of channels and then provides a way to determine how to allocate these channels to best transmit data between the base station 1804 and each of the subscriber units 101, and vice versa. In the illustrated example in Fig. 19, the bandwidth is divided into sixty-four (64) subchannels, each providing an 8 kbps data rate. It should be understood herein that within a CDMA type system, the subchannels are physically implemented by encoding a data transmission with one of a number of different assignable codes. For example, the subchannels may be defined within a single CDMA radio frequency (RF) carrier by using different orthogonal Walsh codes for each defined subchannel. (The subchannels are also referred to as "channels" in the

following discussion, and the two terms are used interchangeably herein).

[0217] As mentioned above, the channels are allocated only as needed. For example, multiple channels are granted during times when a particular subscriber unit 1801 is requesting that large amounts of data be transferred. In this instance and in the preferred embodiment, the single subscriber unit 1801 may be granted as many as 20 of these channels in order to allow data rates of up to 160 kbps (20 * 8 kbps) for this individual subscriber unit 1801. These channels are then released during times when the subscriber unit 1801 is relatively lightly loaded. The invention determines the way in which the limited number of channels are divided at any moment in time among the subscriber units 1801.

[0218] Before discussing how the channels are preferably allocated and deallocated, it will help to understand the general architecture of relevant parts of a typical subscriber unit 1801 and base station 1804 in greater detail. Turning attention now to Fig. 20, the base station 1804 accepts data from incoming data sources 2001 through 2003. Each data source 2001 through 2003 represents any type of data source that is sending data to one or more of the subscriber units 1801. For example, data source 2002 may be web server software on network 1805 serving web pages to a client web browser operating in conjunction with subscriber unit 1801-1, while data source 2003 may be an ISDN terminal on network 1805 that is sending voice and data to subscriber unit 1801-3.

[0219] For each subscriber unit 1801 that is in communication with this particular base station 1804, the base station 1804 establishes and allocates a respective data buffer 2011 through 2013. Data buffers 2011 through 2013 store the data that is to be transmitted to their respective subscriber units 1801. That is, in a preferred embodiment, there is a separate data buffer in the base station 1804 for each respective subscriber unit 1801. As subscriber units enter into and exit out of

communication sessions or connections with base station 1804, the number of buffers may change. There is always a one-to-one correspondence between the number of buffers 2011 through 2013 allocated to the number of subscriber units 1801 communicating with base station 1804. The buffers 2011 through 2013 may be, for example, queues or other memory structures controlled by software, or may be hardware controlled fast cache memory.

[0220] As data is queued up in the buffers 2011 through 2013, transmission processor 2010 transmits the data from the base station 1804 to the respective subscriber units 1801. In the case of forward link transmission (from the base station 1804 to the subscriber units 1801), a selection of a limited number of forward link channels 1810a through 1810c are used. As will be explained, the invention is able to accommodate greater bandwidth for one particular subscriber unit 1801, as more and more data is queued at the base station 1804. That is, as the transmission processor 2010 in the base station 1804 accepts data from each buffer 2011 through 2013 for transmission to that buffers' respective subscriber unit 1801, the transmission processor 2010 uses only the allocated number of forward link 1810 resources assigned to that particular respective subscriber unit. To determine how these channel resources are assigned, the invention provides a channel resource assignor 2009 which implements a unique algorithm according to the invention that monitors buffer usage to determine an urgency characteristic of each subscriber unit 1801 in order to dynamically assign an optimum number of channel resources to be allocated to each subscriber unit.

[0221] In the reverse direction, each subscriber unit 1801 also contains a respective data source 2021 through 2023 that provides data to data buffers 2025 through 2027. The data stored in buffers 2025 through 2027 is data to be transmitted on one or more of the reverse links 1811a-c back to the base station

1804, for eventual transmission to processes or devices on network 1805 that are connected at a network session layer with the subscriber units 1801. Each subscriber unit 1801 also contains a transmission processor 2031 through 2033 for controlling the transmission of data from buffers 2025 through 2027 back to base station 1804. As in the base station 1804, the transmission processors 2031 through 2033 only use an allocated number of reverse channel 1811a-c resources assigned to that particular respective subscriber unit 1801.

[0222] In a preferred embodiment of the invention, the channel resource assignor 2009 in the base station also monitors the usage of buffers 2025 through 2027 within subscriber units 1801. This is accomplished via buffer monitors 2035 through 2037 in each subscriber unit 1801 which periodically report buffer characteristics back to base station 1804. The buffer characteristics reports may be piggybacked onto the regular transmission of data on the reverse links 1811a-c. Upon receipt of this buffer characteristic information, the channel resource assignor 2009 then determines an urgency factor representing the relative need for each subscriber unit 1801 to transmit data on the reverse links 1811a-c from their respective buffers 2025 through 2027. Using these urgency factors, the channel resource assignor 2009 can then dynamically assign an optimum number of channel resources which each subscriber unit may use on the reverse links 1811a-c. This channel assignment information sent back to the subscriber units 1801 on the forward links 1810, so that the transmission processors 2031 through 2033 know their currently allocated channels at all times.

[0223] The channel resource assignor 2009 is thus a bandwidth management function that includes the dynamic management of the bandwidth allocated to a particular network layer session connection. Before a further description of the channel assignor 2009 is given, it should be first understood that no matter what

bandwidth allocation is given to a particular subscriber unit 1801, a network layer communication session will be maintained even though wireless bandwidth initially allocated for transmission is reassigned to other connections when there is no information to transmit. One manner of maintaining network layer communication sessions during periods of reduced allocation of bandwidth for a particular subscriber unit is discussed in detail in the above-referenced co-pending patent applications, the entire contents of which are hereby incorporated by reference in their entirety.

[0224] In general, bandwidth assignments are made for each network layer session based upon measured short term data rate needs as determined by buffer statistics. One or more channels are then assigned based upon these measurements and other parameters such as amount of data in the buffer, the present resources allocated to a subscriber unit, and probabilities of a requirement of a subscriber unit to transmit data or priority of service as assigned by the service provider. In addition, when a given session is idle, a connection is preferably still maintained end to end, although with a minimum number of channel resources allocated, such as a single subchannel being assigned. This single subchannel may eventually be dropped after a predetermined minimum idle time is observed.

[0225] Fig. 21 illustrates a buffer 2160 in detail. Buffer 2160 can be any one of the buffers 2011 through 2013 or 2025 through 2027 in either the subscriber units 1801 or base station 1804. The buffer 2160 accepts data 2165 and stores this data while awaiting transmission on forward links 1810 from the base station 1804 to a respective subscriber unit 1801, or on reverse links 1811 from one of the subscriber units to the base station 1804. Each buffer has associated with it L thresholds, which in this example are labeled 1, 2, . . . L and numbered 2161, 2162 and 2163 respectively. These L thresholds are an indication of how much data is currently

stored in the buffer 2160. That is, the thresholds are "characteristics" in the sense that they provide an indication of how much buffer memory is currently in use.

[0226] As data 2165 enters and fills buffer 2160, until transmission of this data takes place, the data may fill buffer 2160 so much so as to cross certain of the thresholds 2161 through 2163. For instance, in Fig. 21, data blocks 2165-a through 2165-d have just filled buffer 2160 enough to approach the first threshold 2161. The last block of data 2165-n exists between thresholds 2161 and 2162 and so the buffer 2160 has stored data in an amount exceeding the first threshold 2161. In other words, buffer 2160 as shown has a threshold level of "1", corresponding to the first threshold 2161.

[0227] As explained above, the channel resource assignor 2009 in base station 1804 obtains an indication of the threshold level for each buffer 2025 through 2027 in each respective subscriber unit 1801-1 through 1801-3. By determining how much data is in each buffer, the resulting data arrival rates of data to each buffer, and the resources currently allocated to transmit data from a buffer, an urgency factor for each data source attempting to transmit on the reverse links 1811 is computed. A similar computation takes place for each data transmitter on the forward links 1810.

[0228] More particularly, an urgency factor is calculated for each buffer based on these buffer characteristics and indicates the relative need to empty the buffer for that particular receiver as compared to the buffers in other receivers. Given urgency factors for each buffer having data queued for transmission to a waiting receiver, the invention is able to determine how to allocate the available channels to best transmit this data.

[0229] The urgency factor for buffer 2160, for example, is based on statistical information gathered for the accumulation of data 2165. The statistical information

is used to compute probabilities of when data 2165 exceeds or does not exceed certain of the L discrete data thresholds 2161, 2162 and 2163. Thus, as data 2165 enters buffer 2160 and exceeds the first threshold 2161, the urgency factor for that buffer, and hence for the receiver associated with that buffer (i.e., for example, one of the subscriber units 1801 for which data 2165 in buffer 2160 is destined) increases.

[0230] The urgency factor for buffer 2160 is also based upon conditional probabilities of how much time has passed since buffer 2160 has had data 2165 transmitted from the buffer to its intended receiver, as well as how much time has passed since data 2165 has been received at the buffer 2160 for storage until transmission may occur. The urgency factor depends partly on the history of the time that the data level in the buffer exists between each threshold in the buffer and on the number of times each threshold, including the maximum buffer capacity, is exceeded.

[0231] The urgency factor is also based on how close data 2165 is to the last threshold L 2163, which indicates that the buffer is reaching maximum capacity. The urgency factor therefore also accounts for the probability of exceeding the capacity of buffer 2160, based on exceeding the maximum threshold L 2163.

[0232] The channel resource allocator 2009 therefore calculates an urgency factor, U, for each of M buffers, where M is the total number of buffers used in the reverse 1811 and forward 1810 links. The urgency factor for the buffers servicing the forward links 1810 are calculated independently of urgency factors for the other buffers servicing the reverse links 1811, and the buffers servicing each transmission direction of a particular connection between a particular one of the subscriber units 1801 and the base station 1804 are independent of one another.

[0233] At any given time, a given buffer J has a number of channels, N_J,

which is the number of channels already allocated to that particular buffer J. Accordingly, N_J must range from $1<N_J<N_{MAX}$, where N_{MAX} is the maximum number of channel resources that may be assigned to any one particular buffer, and hence to any one link. In the preferred embodiment, N_{MAX} can be as high as 20 channels, with each channel operating at approximately 8.55 kilobits per second (kbps) or at 13.3 kbps, depending upon a rate selection as determined by which CDMA standard is used. Thus, if a particular buffer is assigned the maximum number of channels to accommodate data transfers for high bandwidth applications, instantaneous data rates may be achieved as high as from about 171 kbps to 260 kbps.

[0234] The urgency factor U for a given buffer is equal to the sum of weighted conditional probabilities. Each conditional probability represents the chance of exceeding the last threshold L, within a time frame, T_S, given that the data in the buffer has already exceeded a particular threshold E_i. The time frame T_S corresponds to the maximum time needed to reallocate a resource. The probabilities for an urgency factor U for a single buffer are all computed in a similar manner, but are based upon different thresholds within that buffer. Thus, as the probabilities for each threshold change with the various demands for service, the urgency factor for that particular buffer also changes.

[0235] In a preferred embodiment, the probability of exceeding a particular threshold E_L in time T_S given that another threshold E_i is exceeded is given by:

$$P_{EL}(T_S \mid E_i) = \frac{P_{EL}(E_i) \cdot P_{EL}(T_s)}{P_{EL}(E_j)} \qquad (21)$$

[0236] Threshold E_i is used in the above equation when computing the probability of exceeding a threshold E_L, in a time period T_S, given that the data level in the buffer has already crossed threshold E_j. Since this is an indirect computation, it may be derived from the formula:

$$\frac{\sum(P_{EL}\ within\ T_S\ of\ E_i) / \sum(E_i\ for\ T_S)}{\sum(E_L / E_i)} \tag{22}$$

[0237] The probabilities that make up the urgency factor U for a particular buffer are also weighted before they are summed, such as

$$U = \sum_i P_{E_L}(T_S \mid E_i) \cdot W_i(N) \tag{23}$$

[0238] The weight $W_i(N)$ for each probability is selected to optimize the resource allocation. For example, the weight is selected based upon which threshold is crossed and therefore affects the urgency factor for that buffer by increasing the weight of the summed probabilities used to compute that urgency factor for that buffer.

[0239] Once an urgency factor U for each buffer has been computed, the channel resource assignor 2009 determines how to allocate the available channels among the buffers. This is accomplished in a preferred embodiment by determining which buffer has the highest urgency factor and which one has the lowest. Next, the highest and lowest urgency factors must exceed respective high and low urgency thresholds. If this is true, one resource channel is deallocated from the buffer with the lowest urgency factor and is reallocated to the buffer with the highest urgency factor. In this manner, the channel resources for buffers may change over time based upon the urgency factors of the buffers.

[0240] Also, when N_J is 1, there is only one channel allocated to a particular buffer. In this state, the assigned channel resource may be reallocated (i.e., taken away) to another buffer if there is no data in buffer and if the probability of exceeding the buffer capacity within the time it takes to reassign this initial resource, $P_{E_L}(T_S \mid E_o)$, is less than the probability of reaching the buffer overflow limit $P(E_L)$, which is a predetermined constant.

[0241] Referring to Figures 22-28, these figures describe a dynamic

bandwidth allocation process for multiple access communications using session queues.

[0242] More specifically, Fig. 22 is a block diagram of a system 2200 for providing high speed data service over a wireless connection by seamlessly integrating a wired digital data protocol such as, for example, Transmission Control Protocol/Internet Protocol (TCP/IP) with a digitally modulated wireless service such as Code Division Multiple Access (CDMA).

[0243] The system 2200 consists of two different types of components, including subscriber units 2201-1, 2201-2, . . ., 2201-n (collectively subscribers 2201) as well as one or more base stations 2204 to provide the functions necessary in order to achieve the desired implementation of the invention. The subscriber units 2201 provide wireless data and/or voice services and can connect devices such as, for example, laptop computers, portable computers, personal digital assistants (PDAs) or the like through base station 2204 to a network 2205 which can be a Public Switched Telephone Network (PSTN), a packet switched computer network, or other data network such as the Internet or a private intranet. The base station 2204 may communicate with the network 2205 over any number of different efficient communication protocols such as primary rate ISDN, or other LAPD based protocols such as IS-634 or V5.2, or even TCP/IP if network 2205 is an Ethernet network such as the Internet. The subscriber units 2201 may be mobile in nature and may travel from one location to another while communicating with base station 2204.

[0244] Fig. 22 illustrates one base station 2204 and three mobile subscriber units 2201 by way of example only and for ease of description of the invention. The invention is applicable to systems in which there are typically many more subscriber units 2201 communicating with one or more base stations 2204.

[0245] It is also to be understood by those skilled in the art that Fig. 22 may

be a standard cellular type communication system such as a CDMA, TDMA, GSM or other system in which the radio channels are assigned to carry between the base stations 2204 and subscriber units 2201. This invention, however, applies more particularly to non-voice transmissions, and preferably to digital data transmissions of varying bandwidths. Thus, in a preferred embodiment, Fig. 22 is a CDMA-like system, using code division multiplexing principles for the air interface. However, it is also to be understood that the invention is not limited to using standardized CDMA protocols such as IS-95, or the newer emerging CDMA protocol referred to as IS-95B. The invention is also applicable to other multiple access techniques.

[0246] In order to provide data and voice communications between the subscriber units 2201 and base station 2204, wireless transmission of data over a limited number of radio channel resources is provided via forward communication channels 2210 which carry information from the base station 2204 to the subscriber units 2201, and reverse communication channels 2211 which carry information from the subscriber units 2201 to the base station 2204. The invention provides dynamic bandwidth management of these limited channel resources on an as needed basis for each subscriber unit 2201. It should also be understood that data signals travel bidirectionally across the CDMA radio channels 2210 and 2211, i.e., data signals originating at the subscriber units 2201 are coupled to the network 2205, and data signals received from the network 2205 are coupled to the subscriber units 2201.

[0247] Fig. 23 provides an example of how dynamic allocation of radio bandwidth may take place in system 2200. First, a typical transceiver within a subscriber unit 2201 or the base station 2204 can be tuned on command to any 1.25 MegaHertz (MHz) channel within a much larger bandwidth, such as up to 30 MHz in the case of the radio spectrum allocated to cellular telephony. This bandwidth is

typically made available in the range of from 800 to 900 MHz in the United States. For PCS type wireless systems, a 5 or 10 MHz bandwidth is typically allocated in the range from about 1.8 to 2.0 GigaHertz (GHz). In addition, there are typically two matching bands active simultaneously, separated by a guard band, such as 80 MHz; the two matching bands form a forward and reverse full duplex link between the base station 2204 and the subscriber units 2201.

[0248] Within the subscriber unit 2201 and the base station 2204 transmission processors (i.e., transceivers) are capable of being tuned at any given point in time to a given 1.25 MHz radio frequency channel. It is generally understood that such 1.25 MHz radio frequency carrier provides, at best, a total equivalent of about a 500 to 600 kbps maximum data rate transmission speed within acceptable bit error rate limitations. In the prior art, it was thus generally thought that in order to support an XDSL type connection which may contain information at a rate of 128 kbps that, at best, only about (500 kbps/128 kbps) or only three (3) subscriber units 2201 could be supported at best on each radio channel.

[0249] In contrast to this, the present system 2200 subdivides the available radio channel resources into a relatively large number of subchannels and then provides a way to determine how to allocate these subchannels to best transmit data between the base station 2204 and each of the subscriber units 2201 and vice versa. In the illustrated example in Fig. 23, the bandwidth is allocated to sixty-four (64) subchannels. It should be understood herein that within a CDMA type system, the subchannels are physically implemented by encoding a data transmission with one of a number of different pseudorandom (PN) or orthogonal channel codes. For example, the subchannels may be defined within a single CDMA radio frequency (RF) carrier by using different orthogonal codes for each defined subchannel. (The

subchannels are also referred to as "channels" in the following discussion, and the two terms are used interchangeably from this part onward).

[0250] As mentioned above, the channels are allocated only as needed. For example, multiple channels are granted during times when a particular subscriber unit 2201 is requesting that large amounts of data be transferred. In the preferred embodiment, the single subscriber unit 2201 may be granted as many as 28 of these channels in order to allow data rates of up to about 5 Mega bits per second for an individual subscriber unit 2201. These channels are then released during times when the subscriber unit 2201 is relatively lightly loaded.

[0251] Maximum flexibility can be obtained by adjusting coding rates and modulation types used for each connection, such as the number of channels. One particular scheme for assigning channel codes, Forward Error Correction (FEC) code rate, and symbol modulation types is described in a co-pending U.S. patent application Ser. No. 09/773,253 filed Jan. 31, 2001 entitled "Maximizing Data Rate by Adjusting Code and Coding Rates in CDMA System", given Attorney Docket No. 2479.2021-000 which is assigned to Tantivy Communications, Inc., the same assignee of the present application, which is also hereby incorporated by reference.

[0252] Before discussing how the channels are preferably allocated by the base station 2204 referring to Figs. 22 and 25 the base station 2204 establishes and allocates a respective data buffer 2540-1 through 2540-3. Data buffers 2540-1 through 2540-3 store the data that is to be transmitted their respective subscriber units 2201. That is, in a preferred embodiment, there is a separate data buffer in the base station 2204 for each respective subscriber unit 2201. As subscriber units enter into and exit out of communication sessions or connections with base station 2204 the number of buffers may change. There is always a one-to-one correspondence between the number of buffers 2540-1 through 2540-3 allocated to

the number of subscriber units 2201 communicating with base station 2204. The buffers 2540-1 through 2540-3 may be, for example, queues or other memory structures controlled by software, or may be hardware controlled fast cache memory.

[0253] The particular process which determines how channels are allocated and deallocated may reside in a data services function disposed within the upper layers of the protocols implemented in the base station 2204 and subscriber units 2201.

[0254] Specifically now, referring to Fig. 24, there is shown a protocol layer diagram such as typically associated with third generation (3G) wireless communication services. The protocol layers follow the open system interconnect (OSI) layered model with a physical layer 2220 media access control sub layer 2230 link access control (LAC) sub layer, 2240 and upper communication layers 2250. The physical layer 2220 provides physical layer of processing such as coding and modulation of the individual logical channels. Access to the logical channels is controlled by the various functions in the MAC sub layer 2230 including channel multiplex sub layer 2232 multiplex control channel multiplex sub layer 2231 radio link protocol sub layer 2233 and SRPB 2234. The signaling link access control functionality 22141 is provided in the LAC sub layer 2240.

[0255] Upper layers processing 2250 includes upper layer signaling 2251 data services 2252 and voice services 2253. The particular decision processes to allocate or deallocate channels to particular network layer connections resides therefore in a data services functionality 2252 in the upper layers 2250. The data services functionality 2252 communicates with the radio link protocol 2233 in the MAC sub layer 2230 in order to perform functions such as to send messages to allocate and deallocate channels from end to end as demand requires.

[0256] Turning attention now to Fig. 25, various components of the base station 2204 and subscriber units 2201 will be described now in greater detail in connection with the process for determining when channels should be allocated or deallocated.

[0257] Fig. 25 is a more detailed diagram of the implementation of the session oriented buffering scheme implemented in the data services function 2252. In particular, Fig. 25 shows how this is implemented in the base station 2204. Network layer traffic is routed to the base station 2204 using typical network routing protocols such as Transmission Control Protocol/Internet Protocol (TCP/IP). At the base station 2204 incoming traffic is separated into individual traffic flows destined for separate subscriber units 2201-1, 2201-2, . . . , 2201-n. The traffic flows may be separated such as by examining a destination address field in the TCP/IP header. The individual traffic flows are delivered first to transport modules 2501-1, 2501-2, ..., 2501-n with a transport module 2501 corresponding to each of the intended subscriber units 2201. A given transport module 2501 is the first step in a chain of processing steps that is performed on the data intended for each subscriber unit 2201. This processing chain includes not only the functionality implemented by the transport module 2501 but also a number of session queues 2510, a session multiplexer 2520 and transmission buffers 2540. The outputs of the various transmission buffers 2540-1, 2540-2, ..., 2544-n are then assembled by a transmit processor 2550 that formats the data for transmission over the forward radio links 2210.

[0258] Returning attention now to the top of the Fig. 25 again, each transport module 2501 has the responsibility of either monitoring the traffic flow in such a way that it stores data belonging to different transport layer sessions in specific ones of the session queues 2510 associated with that transport module 2501. For

example, transport module 2501-1 assigned to handle data intended to be routed to subscriber unit 2201-1 has associated with it a number, m, of session queues 2510-1-1, 2510-1-2, . . . , 2510-1-m. In the preferred embodiment, a given session is characterized by a particular transport protocol in use. For example, in a session oriented transport protocol, a session queue 2510 is assigned to each session. Such session transport oriented protocols include, for example, Transmission Control Protocol. In sessionless transport protocols, a session queue 2510 is preferably assigned to each stream. Such sessionless protocols may for example be the User Datagram Protocol (UDP). Thus traffic destined for a particular subscriber unit 2201-1 is not simply routed to the subscriber unit 2201-1. First, traffic of different types are from the perspective of the transport layer are first routed to individual session queues 2510-1-1, 2510-1-2, ... , 2510-1-m, associated with that particular connection.

[0259] Another key function performed by the transport module 2501-1 is to assign priorities to the individual queues 2510-1 associated with it. It will later be understood that depending upon the bandwidth available to a particular subscriber unit 2201 traffic of higher priority will be delivered to the transmission buffer 2540-1 before those of lower priority. This may include traffic that is not session oriented, for example, real time traffic or streaming protocols that may be carrying voice and/or video information.

[0260] More particularly, the transport module 2501-1 reports the priorities of each of the individual session queues 2510-1 to its associated session multiplexer 2520. Traffic of higher priority will be selected by the session multiplexer 2520 for loading into the transmit buffer 2540-1 for loading traffic of lower priority, in general. Traffic of equal priority will either be fairly selected such as using techniques known as weighted fair queuing (WFQ) or other schemes such as oldest

queued data loaded first.

[0261] Priorities associated with each session queue may be obtained from information such as a profile data record kept for each user. For example, some users may have specified that they desire web page traffic traveling on TCP type session connections to have lower priority than streaming audio information carried on UDP type connections. Prioritization may also be based on other aspects of the data content being transmitted. For example, traffic being forwarded from a private data network may be given priority over traffic being forwarded from public networks.

[0262] Each of the session multiplexers 2520-1, 2520-2, ... , 2520-n, reports indications to a session manager 2530 of the states of all of the session queues 2510 that it is currently managing. The session manager 2530 also receives indications of the present forward channel assignments given to each individual subscriber unit 2201 by the channel assigner 2509. The channel assigner 2509 monitors the usage of the transmit buffers 2540 in the base station. Upon receipt of characteristic information concerning the state of how much data is queued in respect to transmit buffers 2540 the channel resource assigner 2509 then determines an urgency factor representing the relative need for each subscriber unit 2201 to receive data on the available forward link radio channels 2210. Using these urgency factors, the channel resource assigner 2509 can then dynamically assign an optimum number of channel resources to be allocated to each subscriber unit 2201. Specific discussion of urgency factors in the allocation of channels is described in further detail below.

[0263] To estimate how much data may be transversing the wired network at any particular instant in time, the session manager 2530 also needs to maintain a running estimate of the latency or the back call network 2205 to any particular server at the other end of a transport layer session. The transport modules 2501

therefore watch individual session flows from various network servers located in the wired network 2205 and are therefore capable of estimating latencies such as by determining typical TCP round-trip time estimations. The transport modules 2501 report this information to the session manager 2530.

[0264] The session manager 2530 containing all of this information can then send channel requests to the channel resource assigner 2509 when it perceives that the present incoming data flow from the wired network for a particular individual subscriber unit 2201-1 is greater than the data rate allowed to that subscriber unit by its present channel configuration. Recalled from above that the channel configuration may include the number of channels assigned, coding rate, and symbol modulation rate for each specific channel. Likewise, the session manager 2530 notifies the channel resource assigner 2509 when it is possible to release channel resources for a particular subscriber unit 2201-1 if the incoming data flow from the wired network 2205 is less than the maximum data rate that is presently assigned to its forward link.

[0265] If split connection transport approaches are employed, as described in RFC 2757-Long Thin Networks, of the Internet Engineering Task Force (EETF), the session manager 2530 is capable of sending requests to the transport modules 2501 that pause data flow for any particular session or sessions. If the session is a TCP session, the transport modules 2501 can then actively place the TCP senders at the other end of the network 2205 into a so-called persist mode, thereby pausing all further session flow. If the session is a streaming or unreliable protocol such as UDP, a loss profile will determine the nature of how the queued and incoming data is lost. Session information will be paused or lost if the session manager 2530 requests that more forward bandwidth should be assigned to a particular subscriber unit 2201-1 and the request denied.

[0266] If channel requests are denied, the session manager 2530 then determines which session information to regulate, pause, or lose data based on content priority information. As previously mentioned, the transport session managers 2501 maintain information to allow them to prioritize their individual session queues 2510 based on content so these transport modules 2501 can therefore choose the correct session queues to enable and/or disable based on priority.

[0267] The transmission buffers 2540 are each marked with levels that are used to calculate urgency factors-for each respective buffer 2540. The urgency factors are used to determine channel allocation by the channel assigner 2509 on a per subscriber per content basis. The levels, indicated in Fig. 25 as L1, L2, and L3, represent demarcation points for channel allocation and/or deallocation. Specifically, when the transmission buffers 2540-1 is filling and a level is traversed, an indication is sent to the channel resource assigner 2509 that the subscriber unit 2201-1 is likely to need more forward link bandwidth assigned. If the request is denied, the channel resource assigner 2509 then sends this indication to the session manager 2530.

[0268] Conversely, when the transmission buffer 2540-1 is emptying, and a level is traversed, an indication is sent to the channel resource assigner 2509 that the associated subscriber unit 2201-1 may have forward traffic channels taken away from or deallocated without affecting end to end performance.

[0269] The levels L1, L2, ... , L3, may therefore be termed under flow thresholds. The levels basically represent permetations of available code rate and channel code assignments for an individual subscriber unit 2201. Two requirements are needed to determine the threshold levels. First, the route trip transfer time on the wired network either needs to be estimated or initial approximation needs to be

set. For TCP sessions, a running round-trip time (RTT) estimation is made. For streaming oriented sessions such as UDP, another approximation can be made which for example may be a function of how much data may be queued to optimize the user's experience for a particular real time application using the UDP protocol.

[0270] Secondly, the data rate over the air interface needs to be determined. This is a function of the present code rate (CR) and number of assigned channels (NCH) allocated to a particular subscriber unit. These are the values determined by the channel resource assigner 2509.

[0271] Coding rates are assigned to subscriber units 2201 determined by the quality of the radio connection. For each assigned coding rate, the subscriber may also be assigned a number of channels. One scheme, therefore, allocates a Level to each available assigned channel. Thus levels L1-LC, where C indicates the number of assigned channels are available at any given instant in time to service the connection. Thus the levels, L1-LC, change each time the number of channels are assigned as well as each time the coding rate changes. Specifically, the particular buffer level associated with each L will change depending upon the available coding rate.

[0272] A graphical representation of a particular transmit buffer 2540 is illustrated in Fig. 26. With knowledge of the round-trip transfer time in the network 2205 and the current available data rate over the forward link radio channels 2210 allocated to the particular subscriber unit 2201 the levels L1-LC may be calculated as follows:

$$Ln = Underflow\ \ Threshold = DR_{Air}(code\ rate\ \&\ channel\ configuration) * \Delta t, \qquad (24)$$

where DR_{Air} is the data rate across the air interface, and the round-trip transfer time is either the estimated time or the set round-trip time over the wired network 105. Δt is the time granularity used to monitor incoming data flows. If this scheme

is used only to optimize TCP connection oriented sessions, Δt can be said to either the maximum or average of all round-trip times estimated by the TCP end points, depending upon the available buffer space.

[0273] The condition for sending a request for more bandwidth to be allocated to a particular subscriber unit 2201 is described by the following relationship:

$$\left[BC_{\Delta t} + \left(\sum_{i=1}^{\max} Fin_i * \Delta t \right) \right] > L(n+1) \qquad (25)$$

where Δt is the time granularity used to monitor the incoming data flows, $BC_{\Delta t}$ represents the current transmission buffer capacity at the beginning of a particular timeframe, Fin_1 minus Fin_{max} represents all incoming data flows from sessions or streams to the transmission buffer 2540 and L(n+1) is the amount of data that can be sent over the radio forward links 2210 in time Δt for the next increasing channel configuration.

[0274] Note that for session oriented TCP streams that the maximum Fin_{Subi} is equal to the maximum advertised received window divided by the round-trip transfer time. This condition occurs when the combination of all incoming flows for a specific time interval is greater than the amount of data that can be transmitted during one time interval Δt at the next increasing channel capacity assignment.

[0275] Fig. 27 represents this case graphically with the block arrow in the Figure representing the amount of flow incoming for the time frame Δt.

[0276] The condition for sending a channel deallocation request for a subscriber unit is given by the relationship:

$$\left[BC_{\Delta t} + \left(\sum_{i=1}^{\max} Fin_i * \Delta t \right) \right] < L(n+1) \qquad (26)$$

where L(n) is the amount of data that can be sent over the assigned forward link

channels 2210 in time Δt for the current channel configuration. This condition occurs when the combination of all incoming flows for a specific time interval, Δt is less than the amount of data that can be transmitted during that time interval at the current channel capacity assignment. This situation is represented in the diagram of Fig. 28 with the block arrow representing the amount of flow incoming during time Δt.

[0277] Note that in an actual implementation, the transmission buffers 2540 may only be theoretical queues represented by a data structure within the session manager 2530 or session multiplexers 2520. The transmission buffers 2540 are actually the combination of all data residing in all session queues 2510 for any particular subscriber unit 2201. This same logic applies when determining urgency factors and levels for the transmission buffer data structures namely that such logic can be implemented within the session manager 2530 and/or session multiplexers 2520 rather than as a separate physical data storage structure and associated logic.

[0278] The present invention therefore provides an advantageous way in which transmission queues may be loaded and how additional resources may be requested and/or may be allocated and/or deallocated on a per subscriber basis. Individual transmission queues intended for particular subscribers may therefore be monitored for data level and channels assigned or deassigned depending upon observed buffer filling rates. The channel resource assigner 2509 therefore has knowledge of the types of traffic flow through the base station based upon application content. This allows more intelligent efficient channel allocation when there is competition for the available resources. Thus by having transport layer aware channel allocation and deallocation coupled with calculation of overflow and underflow threshold based upon current configured forward link radio channel capacity, the connection between the base station and the subscriber unit in the

forward link direction may be optimized.

[0279] While the present invention has been particularly shown and described with references to preferred embodiments thereof, it will be understood by those skilled in the art that various changes in form and details may be made therein without departing from the spirit and scope of the invention as defined by the appended claims. Those skilled in the art will recognize or be able to ascertain using no more than routine experimentation, many equivalents to the specific embodiments of the invention described specifically herein. Such equivalents are intended to be encompassed in the scope of the claims.